MONTHLY FORECASTS BY CORRELATION

JUNE, A KEY MONTH

By CHARLES D. REED

[Weather Bureau Office, Des Moines, Iowa, July 9, 1925]

In Iowa as June mean temperature goes so goes the temperature of the next three months. For the 35 years of State-wide records, ending with 1924, the correlation coefficient of the mean temperature of June with the combined mean temperature of July, August, and September is 0.559 ± 0.078. Expressing this in a regression formula for prediction purposes, it takes the form

$$Y = .34X + 46.37$$

in which X is the mean temperature of the current month of June and Y is the predicted mean temperature of the three following months combined.

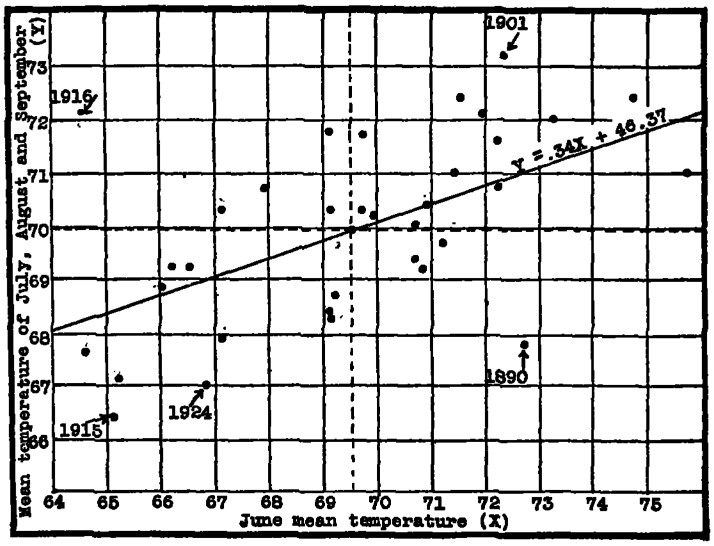


FIG. 1.—June temperature indicates the temperature of the next three months, as shown by the sloping regression or prediction line. Broken lines show mean temperatures. Years that depart most widely connected with dots by arrows

The dispersion of the data can best be visualized by an inspection of Figure 1. It will be observed that there is a fairly well defined arrangement of the data along the regression line. As is usual in such cases, it is much easier and sufficiently accurate to read the prediction from the coordinates instead of computing it by the formula, though the regression line could not, of course, be accurately located in the first place without the necessary mathematical process. Its meaning is that if the formula $Y = .34X + 46.37$ be applied to the prediction of the temperature of each three-month period and the average error determined, the error will be smaller than could be obtained by any other line or formula. Inspection of Figure 1 gives the impression that no curvilinear formula could be devised that would be superior. As a matter of fact, the average error in the 35 cases under consideration is $\pm 1.1°$. The greatest error would have been $-3.8°$ in 1916 and there are two cases, 1915 and 1917, when the error would have been zero.

The error in the temperature predicted by this formula would, in 35 years, have equaled or exceeded 1.5° 10 times, 2.0° only 5 times, 3.0° but twice, and never would have exceeded 3.8°. The probable accuracy is made more apparent in Figure 2, where the actual data, limited to the 35 cases, are expressed in percentage, dotted in, and a smooth curve drawn to indicate the probable accuracy of a large number of cases.

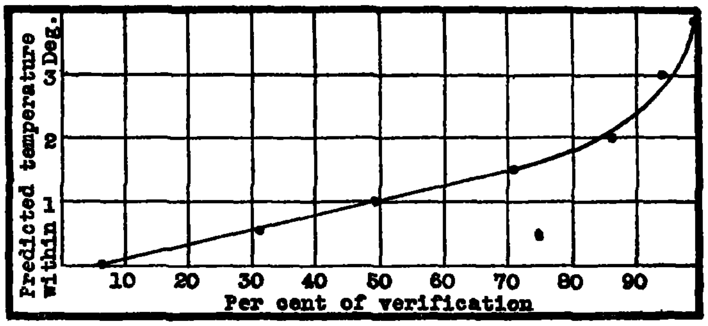


FIG. 2.—Per cent of accuracy for each limiting degree in the prediction; e. g., the prediction will be accurate within 1.5°, 71 per cent of the time

Notable exceptions.—The years that departed most widely from the formula have, for convenient reference, been noted on Figure 1. In the year 1916 some very unusual influence seems to have depressed the temperature in June to the lowest of record and far below the point indicated by correlations with various other meteorological elements. Whatever this influence was, it affected about 90 per cent of the area of the North American continent, and it disappeared abruptly in the closing days of June, so that the following July in Iowa is next to the warmest of record. The mean of the three months, July, August, and September, working the formula backwards, indicates that June should have had a mean of 75.7° instead of 64.5°, the lowest June mean of record in Iowa. The predicted temperature would have been 3.8° too high which is the greatest error in the series. Any influence so strong and with such sharply defined time limits ought to be easily traced. Kimball[3] states that data from American stations showed a noticeable depression in the intensity of solar radiation measured at the earth's surface in 1916–17, and the dots on his graph show a marked decline about June, though more general reports from the Northern Hemisphere received later were not in harmony with the American data.[4] This influence seems to have been peculiar to North America. Volcanic dust could scarcely account for the freak, for it could not be dispersed so rapidly as to permit the following month to become next to the warmest July of record. Nor has this writer been able to find any account of notable volcanic activity near that time.

The next widest departure from the predicted temperature was in 1890, when a warm June was followed by a cool three-month period. In this case, the predicted temperature would have been 3.3° too high.

In the abnormally hot season of 1901, June temperature indicated the hot weather to follow, but the predicted fell short of the actual temperature by 2.2°. As this was the warmest three months (July, August, and September) of record, the formula could not be expected to go all the way. Similarly, the cool June of 1915 indicated the cool season to follow, but lacked 2.1° of going all the way in the coolest three months following.

[3] MONTHLY WEATHER REVIEW, August, 1918, 46: 356.
[4] MONTHLY WEATHER REVIEW, November 1924, 52: 528.

Likewise the cool season of 1924 was indicated by the cool June, but the predicted lacked 2.1° of being as low as the actual. Each of these freak years offers an interesting problem for further study.

Frequency.—While the foregoing presents the quantitative relationship with mathematical accuracy, a simpler and very interesting study was also made of the frequency with which the rule expressed in the opening sentence of this paper was followed. Of the 35 Junes under consideration, 18 were above normal and 17 below normal in temperature. Of the 18 Junes above normal, 14, or 78 per cent, were followed by a mean temperature in the next three months above normal; and of the 17 Junes below normal, 12 or 71 per cent, were followed by a three-month mean below normal. Combining the two, 26 cases out of 35 or 74 per cent followed the rule. This is considerably better than a 50–50 guess and ranks favorably with forecasts that can be made for three or four days in advance. It is possible, however, that a record of 35 years is too short to establish such a rule.

June temperature and July temperature.—The correlation between June temperature and July temperature in Iowa is 0.394 ± 0.096 and the regression formula for predicting July temperature from June temperature is

$$Y = 0.4X + 46.2$$

in which Y is the July temperature required and X is the temperature of the preceding June. The correlation coefficient is not as large as in the case of June temperature with July, August, and September temperatures combined; and the frequency with which June temperature departures are in the same direction as July departures is somewhat smaller. June temperatures were above normal 18 out of 35 times. Ten of these 18, or 56 per cent, were followed by warm Julys. June was below normal 17 times and 12 of the 17, or 71 per cent, were followed by cool Julys. Combining these, 22 out of 35, or 63 per cent, followed the principle that July departures tend to follow June departures.

June temperature indicates July rainfall.—The correlation of June temperature with July rainfall in Iowa is -0.486 ± 0.087. The regression formula is

$$Y = 26.74 - .33X$$

in which Y is the required rainfall of July in inches and X is the mean temperature of the preceding June in degrees Fahrenheit. Though the correlation coefficient is not large, due to a rather wide dispersion of the data (see fig. 3), and the quantitative relationship necessarily shows discrepancies between predicted and actual amounts, dry Julys follow warm Junes with remarkable frequency. Of the 18 Junes with temperatures above normal, 15, or 83 per cent, were followed by deficient rainfall in July, averaging about a third of an inch deficiency for each excess degree of June temperature. When the June temperature was below normal it indicated July rainfall above normal only 9 out of 17 times, or 53 per cent, and is therefore of no value. However, for each degree June temperature is below normal July precipitation averages about a third of an inch above normal.

Here again the June, 1916, temperature seems unaccountably low, and the strange thing is that if the temperature of June, 1916, be taken as 75.7°, as the purely temperature correlations indicated it should have been, as before mentioned, it would have indicated 1.76 inches of rainfall in July following, while the actual amount was 1.78 inches. This is further proof of some temporary, powerful, and unusual influence that depressed surface temperatures far below what they should have been in June, 1916. From what has gone before it is a corrollary that if July is warm it is also dry, and if it is cool it is also wet. The correlation coefficient expressing this fact is -0.502, which harmonizes nicely with the general proposition. For the 18 cases when June temperatures were above normal the formula would have predicted the rainfall of July within 0.5 inch half of the time; within 1.0 inch 72 per cent of the time; within 1.5 inches 83 per cent, etc.

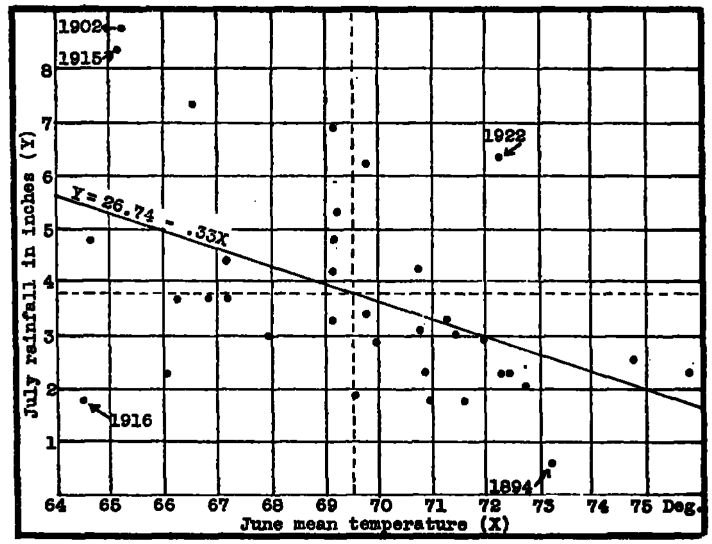


FIG. 3.—The sloping line is the regression or prediction line; broken lines, mean temperature and mean precipitation. A warm June is 83 per cent indicative or a dry July in Iowa, the most notable exception being 1922

Practical form of forecast.—If the June temperature is 2° above normal, the following forecast could be issued for Iowa: July will be drier than the average, by a probability of a little better than 8 chances out of 10. The indicated rainfall is about 3.1 inches, while the average for the last 35 years is about 3.8 inches. The chances that the rainfall will not be greater than 4.1 inches nor less than 2.1 inches are better than 8 out of 10. The chances that July will be warmer than the average are about 6 out of 10; but the temperature will probably be less than a degree above the average.

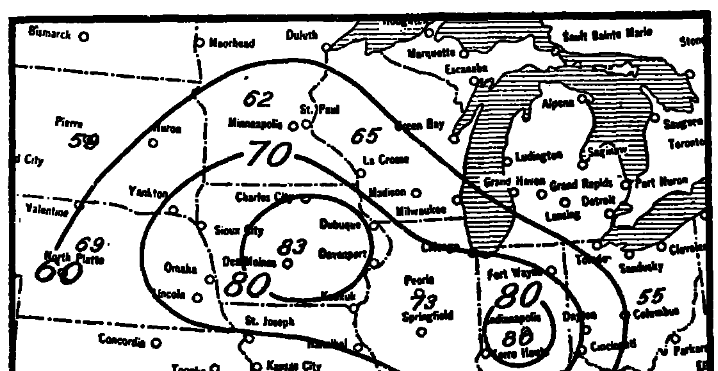
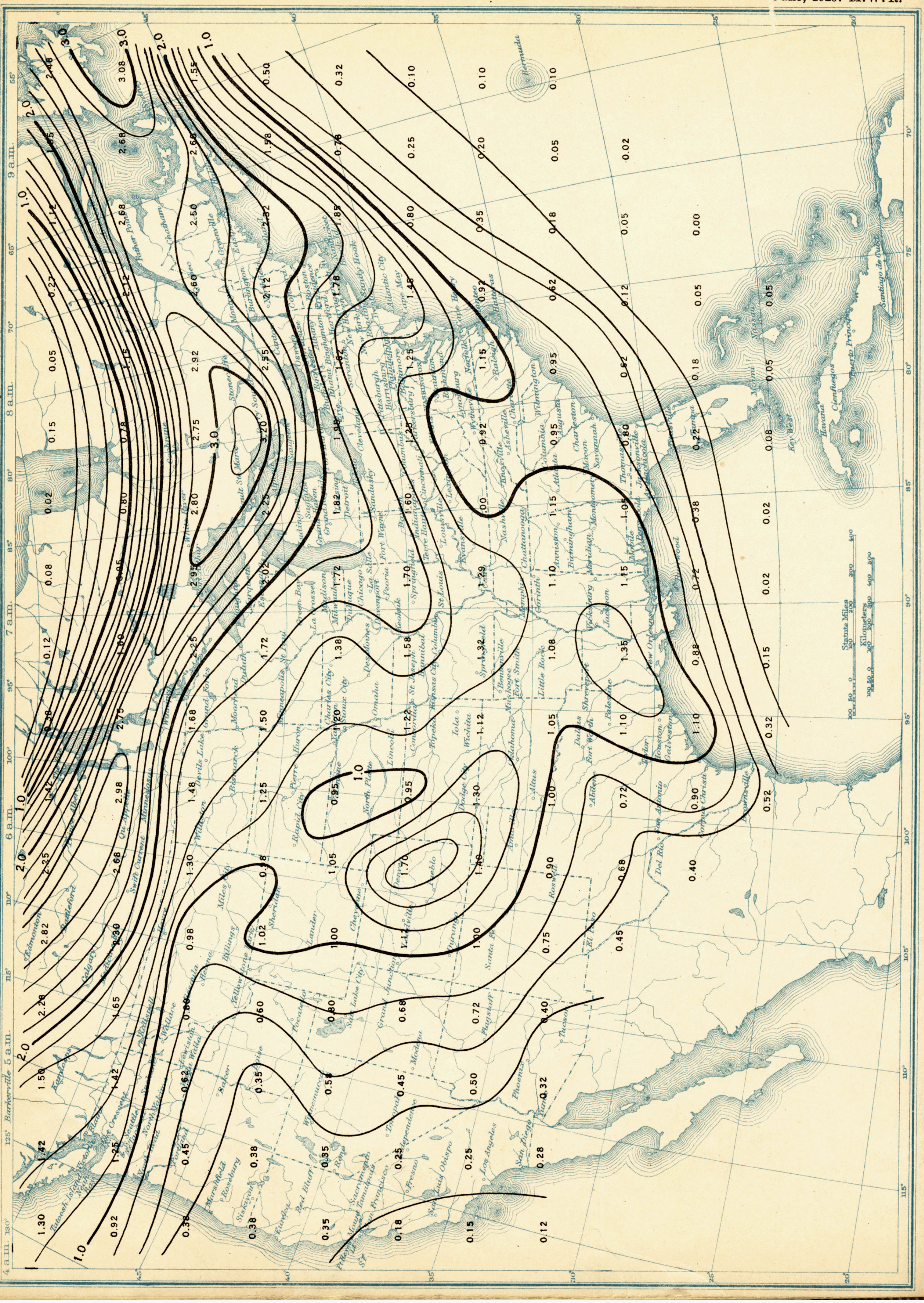


FIG. 4.—Figures near the center of each State show the percentage frequency that dry Julys follow warm Junes

The average temperature of the next three months, July, August, and September, combined will probably be

K. K. **Fig. 2. Isoclones for the United States, January**



above normal, the chances being nearly 8 out of 10. The indicated average of the next three months is about 70.7°, which is 0.7° above normal.

Will this apply elsewhere?—The percentage of frequency that warm Junes are followed by dry Julys has been worked out for near-by States. These percentages appear on the small map, Figure 4. From Iowa to Indiana the chances of verification are better than 7 out of 10, but outside of this the percentage diminishes in regular zones. The verification is somewhat greater by State areas than by individual stations.

Practical applications.—When the end of a cool June has been reached without business activity in warm weather goods, such as palm beach suits, straw hats, bathing suits, electric fans, ice cream, soft drinks, etc,. the merchant wonders if June is a sample of the rest of the season. He could be advised that the chances were better than 7 out of 10 that July and also the next three months would average below normal in temperature. He could than take steps to unload stocks or contract rather than expand his business. His policies and the character of his advertising would be entirely different. Or if June is warm, warm weather enterprises, such as bathing beaches, amusement parks, and water-front real estate, could put on full speed ahead with the assurance that they had nearly 8 chances out of 10 of having a good run of business in the next three months.

Sarle has shown that the merchantability of Iowa corn is more largely determined by June temperature than by any other factor. If June is warm, it is almost safe to assert that very little of the corn will be frosted or immature. There is almost no correlation between merchantability and date of first killing frost in autumn, strange as this may seem, for the real damage is done, or advantage gained, in June.

The relationship between June temperature and corn yield is complicated by so many factors that it can not be expressed by simple correlation. However, the factors can be separated and measured, as this writer hopes to show in a future paper. It is sufficient to say at this time that a warm June produces luxuriant corn plants that are rated at a high percentage condition by crop reporters on July 1. What happens to the corn later on is not a fault in the structure of the plants, but is almost wholly due to the fact that more than 8 times out of 10 a warm June is followed by a dry July, and if dry it is most likely hot, as was shown in the first part of this paper. This cuts the yield but improves the merchantability as a rule.

Other correlations.—No other such large, simple correlations have been found, though nearly 300 have been worked out by computing-machine [5] methods. In a general way a regression formula coming from a correlation smaller than ±0.30 will not give a prediction much better than guess work, yet by the method of partial correlation the indications of several previous months make it possible to say with an accuracy much better than guess work whether or not the month just ahead will be drier, wetter, warmer, or cooler than normal.

[5] "Correlation and Machine Calculation," Wallace and Snedecor, Iowa State College of Agriculture and Mechanic Arts, Official Publication, Vol. 23; No. 35; January 28, 1925.

A NEW METHOD OF CHARTING STORM FREQUENCY

By KEITH KELSEY

[Grand Central Terminal, New York City]

The usual method employed by meteorologists to indicate graphically the storminess of a given area is to plot on a map the number of cyclone centers crossing each 5° square in that area and then to draw lines through points of equal storminess. This method has several disadvantages:

The area of a 5° square in the vicinity of Tampa is, roughly, 30 per cent, or 24,000 square miles greater than the area of one near Duluth.

A chart of isoclones [6] on this basis is thus distorted in favor of the southerly latitudes in the Northern Hemisphere.

"The number of barometric minima per month passing through the 5° square surrounding Duluth" does not convey as clear a conception of conditions as does the more rational and apposite expression, "the number of barometric minima per month passing within 200 miles of Duluth."

A cyclone cutting the corner of the 5° square surrounding Duluth is counted for that city, while one moving due eastward 30 miles closer (*b*, fig. 1) or one passing due southward 85 miles closer (*c*, fig. 1) would not be so counted.

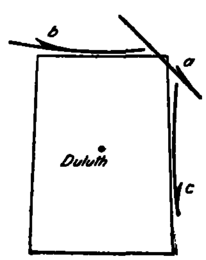


FIG. 1

The 5° square is too large a unit to allow the chart to show to any considerable degree the effect of land and water areas upon cyclone paths.

A unit of area that does not involve these objections is a circle 400 miles in diameter. The shape of this unit gives equal weight to each storm regardless of its direction of travel or its latitude, while the size of the unit is such that sufficient data for good interpolation for an area the size of the United States are obtained.

The accompanying isoclonic chart was prepared using a circle of 400 miles diameter as the unit. The isoclones give the average number of cyclones passing within 200 miles of a given locality during the month of January. The necessary data were obtained from the track charts of the MONTHLY WEATHER REVIEW for the 40 Januarys, 1892 to 1921.

Tangent circles of 200 mile radius were carefully drawn on sheets of onion-skin paper the size of the track charts. Latitude and longitude register marks were drawn on these sheets so that when each was oriented on a track chart the centers of the tangent circles would be in the same geographical localities in every case. Cyclone track charts from 1885 to 1890 were on such a projection as to necessitate the use of ellipses to represent the circles of charts of other years.

One such prepared sheet of onion-skin paper was then clipped onto a cyclone chart with precise adjustment of the register marks. Next taking each storm track in order by number, all the cyclones were followed across the chart, a dot being placed within each circle crossed by the storm path. This same onion-skin sheet was used for 10 Januarys and when all the storm tracks of those 10 months were checked off the dots in each circle were totaled.

Upon the completion of four of these circled charts the figures in corresponding circles were added and the totals divided by 40 giving the average number of paths

[6] Contraction of isocyclone.